

基于双扩张层和旋转框定位的群猪目标检测算法研究

耿艳利^{1,2} 林彦伯¹ 付艳芳³ 杨淑才⁴

(1. 河北工业大学人工智能与数据科学学院, 天津 300130; 2. 智能康复装置与检测技术教育部工程研究中心, 天津 300130;
3. 河北省畜牧总站, 石家庄 050035; 4. 天津魔界客智能科技有限公司, 天津 300130)

摘要: 目前猪群图像检测均为基于水平框的目标检测算法,对于图像中猪体粘连和相互遮挡情况检测率较低,针对图像中的猪只长宽比例较大和可能发生任意角度旋转的特点,提出了一种基于双扩张层和旋转框定位的群猪目标检测算法(Dual dilated layer and rotary box location network, DR-Net)。采集3个猪场的群猪图像,利用数据增强保留9 600幅图像制作数据集;基于膨胀卷积搭建提取图像全局信息的双扩张层,借鉴Res2Net模块改进CSP层融合多尺度特征,猪只目标以旋转框定位并采用五参数表示法在模型训练中利用Gaussian Wasserstein distance计算旋转框的回归损失。试验结果表明,DR-Net对猪只目标识别的精确率、召回率、平均精确率、MAE、RMSE分别为98.57%、97.27%、96.94%、0.21、0.54,其检测效果优于YOLO v5,提高了遮挡与粘连场景下的识别精度和计数精度。利用可视化特征图分析算法在遮挡和粘连场景下能够利用猪只头颈部、背部或尾部特征准确定位目标。该研究有助于智能化猪场建设,可为后续猪只行为识别研究提供参考。

关键词: 群猪; 目标检测; 膨胀卷积; Gaussian Wasserstein distance; 旋转框定位

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2023)04-0323-08 **OSID:** 

Object Detection Algorithm for Pigs Based on Dual Dilated Layer and Rotary Box Location

GENG Yanli^{1,2} LIN Yanbo¹ FU Yanfang³ YANG Shucan⁴

(1. School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China

2. Engineering Research Center of Intelligent Rehabilitation Device and Detection Technology, Ministry of Education, Tianjin 300130, China

3. Hebei Provincial General Animal Husbandry Station, Shijiazhuang 050035, China

4. Tianjin Mojie Technology Co., Ltd., Tianjin 300130, China)

Abstract: At present, the target detection algorithm based on horizontal box is applied to pig objection detection. The adhesion and mutual occlusion in the image of pigs bring great difficulty to individual pig detection. The image of pig has a large ratio of length to width and may rotate at any angle. Object detection algorithm for group pig images based on dual dilated layer and rotary box location network (DR-Net) was proposed. Images of pigs was collected in three pig farms. A dynamic clustering method based on histogram feature and singular value decomposition was used to extract the key frames of pig videos, Laplace operator was used to eliminate images with unclear targets. There were 9 600 images as the data set after data enhancement. The outline of the pig with rotary box was marked. Data set was divided into training set, verification set and test set according to 8: 1: 1. Dual dilated layer used the residual structure and combined two convolution with different dilation factors. The receptive field was increased exponentially with the increase of layers. Stacking dual dilated layers can obtain very large receptive field, it can help the model understand the global information of the image with fewer parameters. Every pig target was located in a rotary box and represented by five parameters. In training, regression loss calculation method based on Gaussian Wasserstein distance was used. The model can get prediction results more accurate. In DR-Net, the features of the input image was extracted by dual dilated layer.

收稿日期: 2022-06-30 修回日期: 2022-09-09

基金项目: 河北省重点研发计划项目(22326606D、20326620D)

作者简介: 耿艳利(1982—),女,副教授,博士,主要从事智能畜牧研究,E-mail: gengyl@hebut.edu.cn

通信作者: 付艳芳(1978—),女,高级兽医师,主要从事生猪饲养管理与环境控制研究,E-mail: fyt781018@163.com

The CSP layer containing multi-layer Res2Net module, which was used to feature fusion and feature extraction of different scales. The prediction results were output through head network. The results showed that the precision, recall, mean average precision, MAE and RMSE of DR-Net were 98.57%, 97.27%, 96.94%, 0.21 and 0.54, respectively. DR-Net was superior to YOLO v5 and YOLO v5 with rotary box location and pig target recognition accuracy was improved. By analyzing the visualization feature map, DR-Net can accurately locate the target using the head, neck, back or tail feature of pigs under occlusion and adhesion condition. The research can contribute to the construction of intelligent pig farm and provide reference for the subsequent research on pig behavior recognition.

Key words: pigs; object detection; dilated convolution; Gaussian Wasserstein distance; rotary box location

0 引言

中国对于猪肉制品的需求量在逐年增加^[1],生猪养殖密度也在不断增加。为了疾病防控和猪场安全生产,在猪舍内安装监控系统有助于管理者监督猪场的整个生产过程,观察监督猪场的生产情况,尽可能地减少进入养殖场的人数,以避免人畜之间的交叉感染^[2]。人为观察和分析监控系统有着极强的主观性,且浪费了大量的时间。基于监控系统的人工智能技术可以对监控画面自主进行分析,但图像中猪只粘连和相互遮挡给检测带来较大困难,群猪目标检测成为重要研究方向。

在畜牧业领域,通过现场采集图像进行猪只识别和行为检测,对生猪养殖业的规模化发展有着重要意义。近几年,已有学者将 Mask R-CNN^[3]、YOLO v3^[4]、SSD^[5]、Faster R-CNN^[6]应用于猪只个体检测^[7-8]。文献[9-16]中的目标检测算法采用水平包围框表示目标,用此类方法对猪只进行检测时存在如下问题:①水平包围框与目标真实轮廓贴合度较差,无法准确反映其对空间的占有情况。②当数个狭长的目标以倾斜的角度密集排列时水平包围框彼此重叠,高度重叠的包围框更容易产生混淆,在后续处理环节中会相互抑制,造成漏检。③目标倾斜时采用的水平包围框中会包含较多的背景信息,对训练和检测带来困难。利用旋转框替代水平框的表示方法,设计基于旋转框的检测模型,用更加紧凑的包围框来进行方向性目标的检测,对识别能力的提升有着重要的价值。

目前已有基于旋转框的目标检测算法应用于遥感影像的研究^[17-19]。本文以提升相互粘连和遮挡场景下的群猪识别精度和减少参数量为目标,将猪只目标以旋转框表示,采用膨胀卷积对图像进行全局特征提取,利用基于 Gaussian Wasserstein distance 的旋转框回归损失计算方法^[20]训练模型,提出一种基于双扩张层和旋转框定位的群猪目标检测算法。

1 群猪图像数据集建立

本文所研究的生猪图像分别采集于天津蓟州区某养猪场、河北省石家庄市某养猪场以及河北省保定市某养猪场。如图1所示,3个猪场群养栏面积和饲养密度有较大差异,其中蓟州猪场为室外环境,群养栏面积为4 m×2.7 m,单栏内平均饲养10头猪,饲养密度为0.98头/m²;石家庄猪场为室内环境,群养栏面积为7.5 m×4 m,单栏内平均饲养10头猪,饲养密度为3.0头/m²;保定猪场为室内环境,群养栏面积为8.4 m×6.5 m,单栏内平均饲养33头猪,饲养密度为1.65头/m²。本文采用 Jetson Nano 开发板和手持设备对不同场景的生猪进行视频采集,拍摄得到的视频总时长为500 min。所有视频均为自然光照条件采集,采集时间为10:00—17:00。



图1 养猪场群养栏环境

Fig.1 Pig farm group enclosure environments

为防止数据重复冗余,对视频进行关键帧^[21]提取,建立群猪图像数据集。关键帧是猪的姿态发生变化的图像,通常包含更多的运动信息,其余图像为正常帧。基于直方图特征和奇异值分解的动态聚类方法是一种有效的关键帧提取方法。图像直方图特征提取方法如图2所示,将每帧图像分成3×3块,每个块生成RGB三通道的直方图,其中每个通道有6个bin,所以每个块中包含216(6×6×6)维特征值,然后将9个直方图串联起来,为每一帧形成一个1944(3×3×6×6×6)维的特征向量。整个视频创

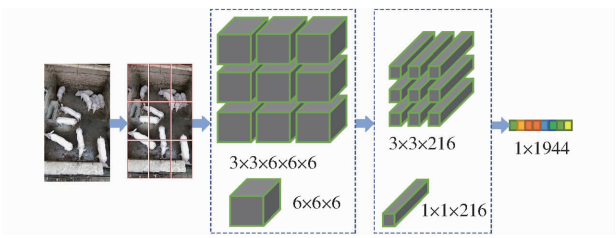


图 2 图像直方图特征提取方法

Fig. 2 Image histogram feature extraction method

建一个维度为 $F \times 1\,944$ 的特征帧矩阵, F 表示视频中的帧数。使用奇异值分解对矩阵进行降维操作, 降维后的特征帧矩阵为 $F \times 63$ 维。最后使用动态聚类方法, 该方法以余弦相似度来创建连续帧的簇, 检查新帧是否与最后形成的簇相似, 选择簇中最后添加的帧作为关键帧。

图像中存在运动模糊和目标重影现象, 利用拉普拉斯算子进行图像预处理剔除目标不清晰的图像。将 RGB 图像转换为灰度图像后, 再经拉普拉斯算子卷积运算后得到图像梯度矩阵, 该矩阵的方差定义为图像清晰度。保留 3 200 幅图像作为数据集, 其中包括蓟州猪场 1 200 幅、石家庄猪场 1 000 幅、保定猪场 1 000 幅。使用 LabelImg2 图像标注工具以旋转框标注出猪只外轮廓, 每幅图像的标注文件为 XML 格式。上述数据集采用水平翻转、垂直翻转和随机旋转等数据增强技术后得到新的数据集共计 9 600 幅, 按 8:1:1 划分为训练集、验证集和测试集, 每个集合中 3 个试验猪场的图像样本比例均为 6:5:5。

2 基于双扩张层和旋转框的群猪目标检测算法

一阶段目标检测算法由主干网络、特征融合网络和预测网络组成, 主干网络负责生成图像不同感受野下的特征信息, 由特征融合网络将各种尺度的特征进行双向融合, 最后由堆叠了 1×1 和 3×3 卷积层的预测网络得到最后输出结果。而两阶段的目标检测算法更加复杂, 在特征融合网络之后首先经过区域建议网络, 生成大量可能存在目标的感兴趣区域, 将这些区域调整成相同尺寸后通过堆叠了全连接层和卷积层的预测网络生成输出结果^[22]。

2.1 双扩张层

双扩张层^[23] (Dual dilated layer, DDL) 中采用残差结构^[24]并结合了两个具有不同膨胀因子的卷积。两个并行的不同膨胀因子的卷积得到与输入尺寸相同的特征图, 然后将两者拼接后用 ReLU 函数激活, 再经过核为 1×1 的卷积层恢复成输入尺寸。其中一个并行的卷积在较低的层中具有较低的膨胀因子, 并且随着层数的增加而呈指数增加。而另一

个并行卷积, 从较低层的大膨胀因子开始, 并随着层数的增加而呈指数减小。感受野随着层数的增加呈指数增长, 堆叠多个双扩张层可以获得非常大的感受野, 这有助于模型用较少的参数量理解图像的全局信息。用于融合多尺度特征的特征金字塔网络在目标检测算法中取得了成功, 但它的多尺度特征是通过最大池化操作获得的, 这些池化操作会导致目标检测所必需的细粒度信息丢失。相反, DDL 既结合了多尺度特征还保持了图像的分辨率。双扩张层结构如图 3 所示。

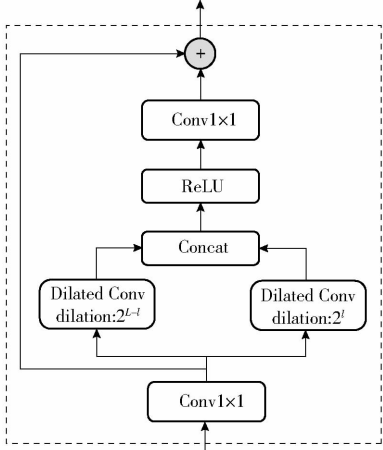


图 3 双扩张层结构图

Fig. 3 Structure diagram of double dilated layer

2.2 Res2Net 模块

特征提取网络通过堆叠卷积层使得网络由细粒度到粗粒度地学习多尺度特征。如图 4 所示, Res2Net 是在 Bottleneck 模块基础上进行改进的模块, 其结构简单并且能够有效地融合多尺度特征。Res2Net 模块整体采用残差结构, 在核为 1×1 的卷积层后, 将特征图分为 s 个子集, 用 G_i 表示, 其中 $i \in \{1, 2, \dots, s\}$ 。每个子集的通道数是原来的 $1/s$, 每个特征图子集 G_i 都有着和原始特征图集相同的空间大小。除 G_1 外, 每个特征图子集 G_i 都有其对应的 3×3 卷积

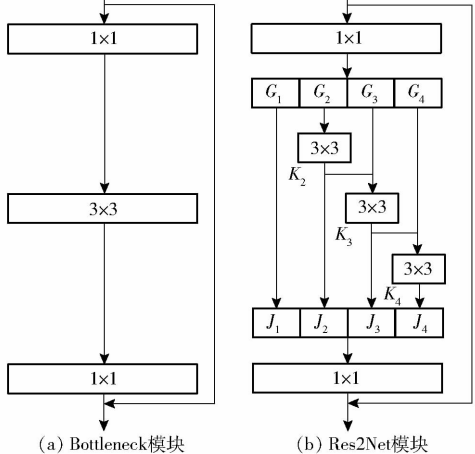


图 4 Bottleneck 模块和 Res2Net 模块对比

Fig. 4 Comparison of Bottleneck module with Res2Net module

层,用 $K_i()$ 表示,其输出为 J_i 。特征图子集 G_i 和 $K_{i-1}(G_{i-1})$ 相加后被一同输入 $K_i()$ 进行处理。每一次特征图子集 G_i 通过一个 3×3 卷积层之后,输出结果就可以获得更大的感受野。 J_i 并联在一起然后通过一个 1×1 卷积层进行信息融合。这种分组、合并的策略使得卷积层能够更有效地处理特征图。

2.3 旋转框回归损失计算

旋转框采用五参数表示法,包括目标中心点位置坐标、目标宽度、高度和角度5个参数表征旋转矩形信息,其中角度周期为 180° ,角度范围为 $[-90^\circ, 90^\circ)$,如图5所示。边界值 -90° 和 90° 实际重合,但数值上有 180° 的差距产生了突变,存在边界不连续问题。

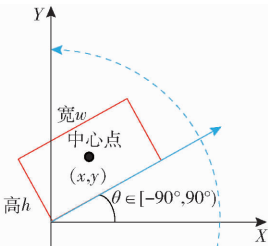


图5 五参数表示法

Fig. 5 Five parameter representation method

旋转框交并比 (Intersection over union, IOU) 不可微分,其回归损失计算由 Smooth-L1 主导^[25],但是 Smooth-L1 得到损失值和评估模型检测精度的 IOU 存在不一致问题。基于 GWD (Gaussian Wasserstein distance) 的回归损失可作为解决上述两种问题的基本方法,将旋转框转换为二维高斯分布近似计算不可微分的旋转 IOU 损失,即使两个旋转边界框之间没有重叠,GWD 仍然可以衡量边界框的损失值。任意旋转框 $B(x, y, w, h, \theta)$ 转换成二维高斯分布 $N(\mu, \Sigma)$,其中 μ 表示二维高斯分布均值,它是由旋转框的中心点横坐标 x 和纵坐标 y 组成, Σ 计算式为

$$\Sigma = \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}} \tag{1}$$

其中

$$\Sigma^{\frac{1}{2}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \frac{w}{2} & 0 \\ 0 & \frac{h}{2} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \frac{w}{2}\cos^2\theta + \frac{h}{2}\sin^2\theta & \frac{w-h}{2}\cos\theta\sin\theta \\ \frac{w-h}{2}\cos\theta\sin\theta & \frac{h}{2}\cos^2\theta + \frac{w}{2}\sin^2\theta \end{bmatrix} \tag{2}$$

式中 θ ——旋转框角度
 w ——旋转框宽度
 h ——旋转框高度
 Σ ——二维高斯分布的协方差矩阵
 $N_1 = (\mu_1, \Sigma_1)$ 和 $N_2 = (\mu_2, \Sigma_2)$ 分别表示两个高斯分布,它们之间的 Wasserstein 距离 d 的计算公式为

$$d = \|\mu_1 - \mu_2\|_2^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{\frac{1}{2}}\Sigma_2\Sigma_1^{\frac{1}{2}})^{\frac{1}{2}}) \tag{3}$$

式中 μ_1 ——二维高斯分布 N_1 的均值
 μ_2 ——二维高斯分布 N_2 的均值
 Σ_1 ——二维高斯分布 N_1 的协方差矩阵
 Σ_2 ——二维高斯分布 N_2 的协方差矩阵
 Tr ——矩阵的对角线元素之和

单独的 GWD 可能对大损失很敏感,因此需要对 d 进行非线性变换使损失更加平滑和富有表现力,旋转框回归损失函数计算式为

$$L_{gwd} = 1 - \frac{1}{1 + \ln(1 + d)} \tag{4}$$

式中 L_{gwd} ——旋转框回归损失值

2.4 DR-Net

针对群猪图像特点,结合上述特征提取方法和旋转框 IOU 损失函数计算方法,提出一种基于双扩张层和旋转框定位的群猪目标检测算法 DR-Net。首先双扩张层对输入图像提取大感受野和小感受野的特征,然后降低图像分辨率再次用双扩张层提取特征,接着利用包含多层 Res2Net 模块的 CSP 层进行特征融合并提取不同尺度的特征,最后经过 Head 网络输出预测结果。Head 网络输出形式与 YOLO v5 相同,输出 3 种尺寸结果。每种尺寸包含 3 种尺寸的水平预选框和 6 种角度参数,共组合成 18 种预选框,每个旋转预选框包含 7 个微调参数:置信度、类别概率、中心点横轴坐标、中心点纵轴坐标、检测框宽度、检测框高度和角度。在模型训练中,采用 GWD 方法计算预测框与真实框之间的损失值,DR-Net 网络结构如图 6 所示。

3 算法运行环境及评价指标

试验训练环境为 Ubuntu 16.04,硬件配置 CPU 为 E5-2680 4 核,主频 2.60 GHz,内存为 32 GB, GPU 为 11 GB NVIDIA GeForce GTX 1080 Ti,试验代码采用 PyTorch 1.7.0 框架进行模型编写。

为验证所提猪只检测算法的有效性,本文采用精确率 (Precision, P)、召回率 (Recall, R)、平均精确率 (Average precision, AP)、F1 值 (F1 score)、

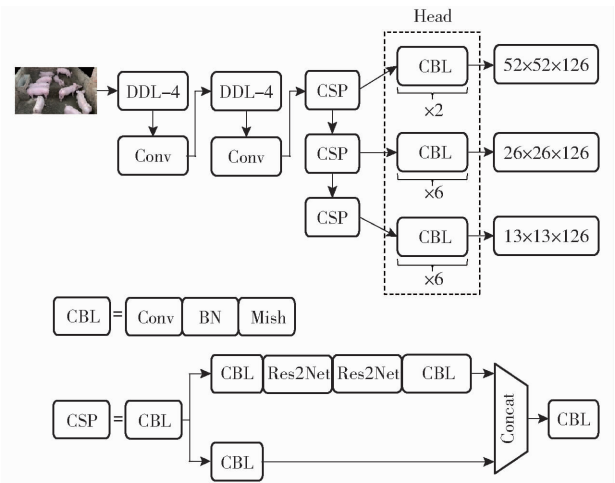


Fig. 6 Network structure diagram of DR-Net

参数量 (Total parameters, Params)、检测速度 (Frames per second, FPS)、模型内存占用量作为评价指标。

深度学习的计数相关研究通常使用平均绝对误差 (Mean absolute error, MAE) 和均方根误差 (Root mean square error, RMSE) 评价模型的有效性。MAE 为预测值与真实值的平均误差,表示算法准确性,MAE 越小,算法准确性越高;RMSE 为预测值与真实值误差的分散程度,表示算法稳定性,RMSE 越小,算法鲁棒性越高。

表 1 DR-Net 的主干网络参数设置和评价指标

模型	DDL 层数		CSP 层中 Res2Net 模块数			平均精确率/%	参数量	检测速度/ (f·s ⁻¹)
	第 1 阶段	第 2 阶段	第 1 阶段	第 2 阶段	第 3 阶段			
DR-Net 1	3	3	2	2	2	95.47	2.46 × 10 ⁷	17.2
DR-Net 2	3	3	4	4	4	95.61	2.67 × 10 ⁷	16.1
DR-Net 3	3	3	6	6	6	95.68	2.88 × 10 ⁷	15.2
DR-Net 4	4	4	2	2	2	96.94	2.47 × 10 ⁷	14.0
DR-Net 5	4	4	4	4	4	96.75	2.68 × 10 ⁷	13.3
DR-Net 6	4	4	6	6	6	96.86	2.89 × 10 ⁷	12.6
DR-Net 7	5	5	2	2	2	96.89	2.48 × 10 ⁷	11.5
DR-Net 8	5	5	4	4	4	96.77	2.69 × 10 ⁷	11.1
DR-Net 9	5	5	6	6	6	96.94	2.90 × 10 ⁷	10.5

4.2 不同损失函数对性能的影响

为了验证采用 GWD 计算旋转框损失能够提升猪只检测精度,采用改进 YOLO v5 作为旋转框目标检测模型,输出与 DR-Net 格式相同。将 Smooth-L1 和 GWD 两种方法在 DR-Net 和改进 YOLO v5 模型上进行性能对比。从表 2 中可看出,改进 YOLO v5 与 DR-Net 在采用 GWD 回归损失计算方法后,模型的精确率、召回率、F1 值、平均精确率分别提高 0.31、0.42、0.35、1.05 个百分点和 1.29、1.37、1.33、1.07 个百分点,模型的 MAE

4 试验结果分析

4.1 DR-Net 模型参数设置

为了得到较优的群猪检测的网络参数和模型,本文分别对 DR-Net 主干网络的双扩张层的第 1、2 阶段层数和 CSP 层第 1、2、3 阶段中 Res2Net 模块串联个数进行调整。调整后的网络结构模型的参数设置、平均精确率、参数量和检测速度如表 1 所示。网络模型的参数对平均精确率和检测速度有明显的影响。从最终检测的平均精确率来看,模型 4~9 的检测效果相近且优于模型 1~3,在检测速度方面模型 4~6 优于模型 7~9,因此双扩张层的第 1、2 阶段层数均为 4。模型 1、4、7 与具有相同双扩张层层数的模型在精确率上相近,模型 3 较模型 1 的平均精确率高 0.21 个百分点,但检测速度低 2 f/s。模型 4 与模型 9 的平均精确率达到 96.94%,但模型 4 比模型 9 在检测速度方面提高 3.5 f/s,因此 CSP 层第 1、2、3 阶段中 Res2Net 模块串联个数均为 2。从整体来看,双扩张层的层数对算法目标检测准确率影响较大,Res2Net 模块串联个数对准确率影响较小。本文采用模型 4 为群猪目标检测模型,即确定 DR-Net 模型双扩张层数为 4,Res2Net 模块串联个数为 2。

和 RMSE 分别降低 0.22、0.31 和 0.21、0.32。可以看出 GWD 能够提高目标检测精度和减小计数误差,因此验证了旋转框的 GWD 回归损失计算方法的有效性。

4.3 DR-Net 性能综合分析

比较 DR-Net 与 YOLO v5 分别在水平框和旋转框下的目标检测和计数性能,如表 3 所示。在本数据集中,两种模型在采用旋转框的模型预测时的平均精确率和计数精度明显优于采用水平框的模型。在水平框下 DR-Net 的精确率、F1 值、平均精

表 2 旋转框两种回归损失计算方法对比

Tab.2 Comparison of two regression loss calculation methods for rotary box

模型	回归损失计算方法	精确率/%	召回率/%	F1 值/%	平均精确率/%	MAE	RMSE
YOLO v5	Smooth - L1	96.62	92.95	94.74	91.10	0.68	1.19
	GWD	96.93	93.37	95.09	92.15	0.46	0.88
DR - Net	Smooth - L1	97.28	95.90	96.58	95.87	0.42	0.86
	GWD	98.57	97.27	97.91	96.94	0.21	0.54

表 3 水平框和旋转框下 DR - Net 与 YOLO v5 的性能比较

Tab.3 Performance comparison of DR - Net and YOLO v5 in horizontal and rotary box

模型	锚框	精确率/ %	召回率/ %	F1 值/%	平均精确 率/%	参数量	MAE	RMSE	检测速度/ (f·s ⁻¹)
YOLO v5	水平	94.35	92.31	93.31	89.65	5.16 × 10 ⁷	0.75	1.27	26.3
DR - Net	水平	94.56	92.30	93.41	90.01	2.45 × 10 ⁷	0.57	0.98	14.0
YOLO v5	旋转	96.93	93.37	95.09	92.15	5.18 × 10 ⁷	0.46	0.88	26.3
DR - Net	旋转	98.57	97.27	97.91	96.94	2.47 × 10 ⁷	0.21	0.54	14.0

确率均略高于 YOLO v5,且 MAE 比 YOLO v5 降低 0.18;而在旋转框下 DR - Net 的精确率、召回率、平均精确率相较于 YOLO v5 提高 1.64、3.9、2.82、4.79 个百分点,在 MAE 和 RMSE 方面相较于 YOLO v5 降低 0.25、0.34。在相同模型中输出旋转框需要比输出水平框多一个角度参数,因此模型参数量增加 2×10^5 ,对检测速度没有影响。DR - Net 的模型参数量只有 YOLO v5 的 1/2 左右,但检测速度降低了一倍,其原因是在高分辨率的图像上为了减少信息损失没有池化操作,而使用多层膨胀卷积,降低了检测速度,但有利于提高模型对全局信息的理解。

分析上述现象,在旋转框中包含较少的背景信息,从而降低了模型对学习能力的要求,使 DR - Net 和 YOLO v5 模型能够得到更精确的定位结果。DR - Net 在水平框和旋转框下均表现出良好的精确率和更低的计数误差,说明双扩张层对图像全局信息具有更强的学习能力,因此基于旋转框的 DR - Net 更适合用于群猪目标检测。

图 7 为 3 个猪场所采集的原始图像及模型检测结果,上图为原始图像,下图为检测结果。表 4 为 960 幅测试集图像的精度统计结果,其中石家庄和保定猪场分别测试 300 幅图像,蓟州猪场测试 360 幅图像。蓟州猪场与石家庄猪场的平均精确率、平均绝对误差和均方根误差相近。蓟州猪场是室外环境,其光照条件较好,图像样本清晰度最高且猪只轮廓明显,有利于猪只个体识别。石家庄猪场的单栏平均猪只数量与蓟州猪场相同,但其栏内面积大、猪只密度低且多处于分离状态,在该场景下平均精确率达到 97.49%,优于其它场景,说明 DR - Net 在低密度场景中识别效果最优。保定猪场的单

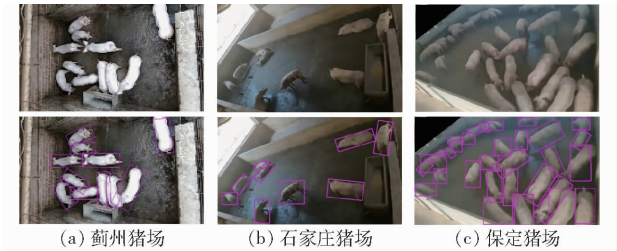


图 7 DR - Net 在不同猪场的检测结果

Fig.7 Detection results of DR - Net in different pig farms

表 4 DR - Net 在不同猪场中的测试结果

Tab.4 Test results of DR - Net in different pig farms

猪场	单栏生猪 平均头数	精确 率/%	召回 率/%	F1 值/%	平均精 确率/%	MAE	RMSE
蓟州	10	98.13	97.71	97.91	97.23	0.18	0.47
石家庄	10	98.53	98.27	98.39	97.49	0.15	0.44
保定	33	96.59	96.30	96.44	95.94	0.28	0.65

栏生猪数量较多、猪只密度大,猪只相互遮挡和粘连的情况较多,虽然识别难度较高但其平均精确率依然达到 95.94%。

4.4 遮挡和粘连场景下检测结果分析

利用 Eigen - CAM^[26]对 DR - Net 模型的检测效果进行特征可视化分析,可以清楚地显示网络在进行图像检测时重点关注的图像区域。图 8 为 DR - Net 在遮挡区域和粘连区域的可视化特征图和检测结果。原图上有 3 种颜色矩形框,红色实线框表示遮挡区域,绿色短划线框表示粘连区域,蓝色点状线框表示漏检区域。

可视化特征图中颜色越红说明网络提取的特征覆盖到该区域越多。遮挡区域是猪只目标被其它物体遮挡造成轮廓不完整或不连续,粘连区域是多个猪只目标姿态相似且相互靠近、轮廓变形和接触面积较大。图像 1、2、3、5 的遮挡区域中模型能够区分

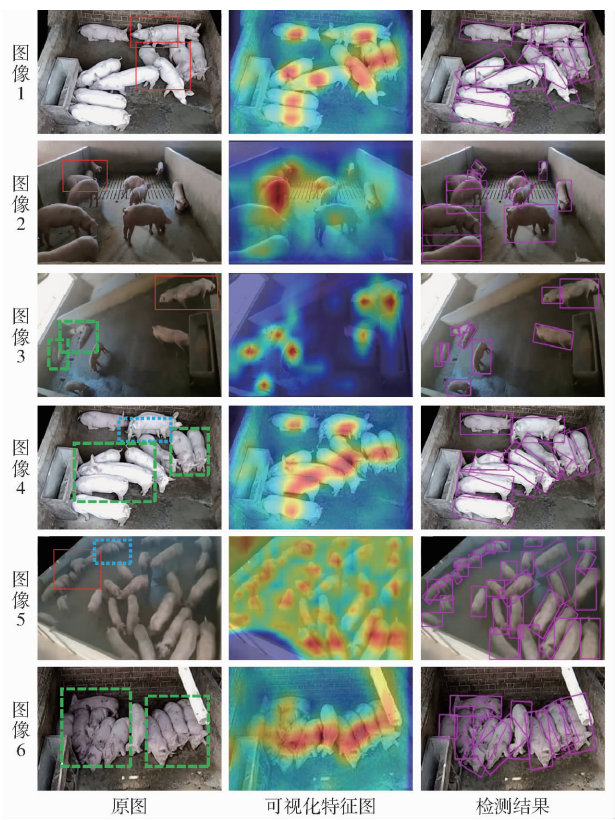


图 8 DR-Net 的可视化特征图与检测结果

Fig.8 Feature visualization and detection results of DR-Net

被遮挡的目标,在可视化特征图上能够看出模型可以从猪只的头颈部、尾部或背部信息推测出被遮挡

猪只所在位置。图像 4 的漏检区域中被遮挡生猪仅有腹部与少部分肢体露出,模型从该处识别猪只位置的能力较弱,原因是数据集中此类高度遮挡的数据较少,并且猪只的四肢占像素面积较少,模型对其关注权重较低。图像 5 的漏检区域中出现多头猪只扎堆现象,猪只之间既有粘连也有遮挡并且像素面积小,给识别带了较大困难。DR-Net 中虽然省去了一阶段目标检测模型常用的特征融合网络,但是依然能学习到猪只的有效特征,说明在高分辨率下利用双扩张层提取全局信息可以达到特征融合网络的效果。图像 3、4、6 的粘连区域中模型的识别效果较好,说明旋转框定位有助于提高粘连场景下的群猪检测精度,原因是训练中以旋转框标记的目标其背景信息少使得猪只特征更清晰,模型输出的预测框带有猪只的角度信息,降低了粘连猪只预测框相互抑制的可能性。

将本文提出的网络 DR-Net 与 YOLO v5 的检测效果进行对比,如图 9 所示。试验结果表明,DR-Net 输出的预测框与目标真实轮廓贴合度较好,能更准确地反映空间占用情况。当多个狭长的目标以倾斜的角度密集排列时,DR-Net 检测效果更好并且不易造成漏检现象。该研究有助于智能化猪场的建设,可为后续猪只行为识别提供参考。

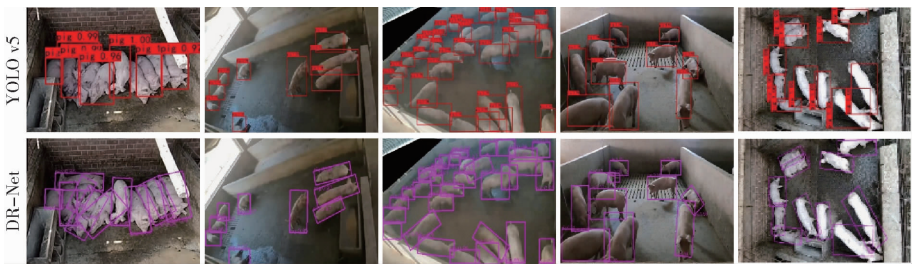


图 9 DR-Net 与 YOLO v5 的检测结果对比

Fig.9 Comparison of detection result between DR-Net and YOLO v5

5 结论

(1)提出了基于双扩张层和旋转框定位的群猪目标检测算法,使用双扩张层在原始分辨率下提取全局特征,利用改进 CSP 层在主干网络中对不同尺度特征进行融合与提取,并采用旋转框提高定位的准确度。模型训练时旋转框的回归损失采用 GWD 方法。该算法提高了群猪目标检测精度,使 AP 达到 96.94%,MAE 降低至 0.21。

(2)验证了旋转框的 GWD 回归损失计算方法的有效性。将 GWD 与 Smooth-L1 方法在改进

YOLO v5 和 DR-Net 上进行性能对比,结果表明,GWD 方法使改进的 YOLO v5 在平均精确率方面提高 1.05 个百分点,MAE 和 RMSE 分别降低 0.22、0.31;GWD 方法使 DR-Net 在平均精确率方面提高 1.07 个百分点,MAE 和 RMSE 分别降低 0.21、0.32。

(3)通过特征可视化分析发现,DR-Net 将注意力集中在猪只的头颈部、背部或尾部,以此来识别遮挡和粘连场景下的猪只个体。双扩张层和旋转框定位可以简化模型结构并有助于提高群猪目标识别的精确率。

参 考 文 献

[1] 黄辉龙,王颖,刘友超,等. 2021 年中国猪肉进出口市场回顾及 2022 年展望[J]. 中国猪业, 2022, 17(1): 38-49.

- [2] 徐占武, 高梨莉. 浅谈视频监控系统的研究现状与应用发展[J]. 盐业与化工, 2013, 42(8): 6-8.
- [3] HE K M, GKIOXARJ G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.
- [4] REDMON J, FARHADI A. Yolov3: an incremental improvement[J]. arXiv preprint arXiv: 1804. 02767, 2018.
- [5] LIU W, ANGUELOY D, ERHAN D, et al. SSD: single shot multiBox detector[C] // European Conference on Computer Vision, 2016: 21-37.
- [6] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C] // Advances in Neural Information Processing Systems, 2015: 91-99.
- [7] 李丹, 张凯锋, 李行健, 等. 基于 Mask R-CNN 的猪只爬跨行为识别[J]. 农业机械学报, 2019, 50(增刊): 261-266, 275. LI Dan, ZHANG Kaifeng, LI Xingjian, et al. Mounting behavior recognition for pigs based on Mask R-CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 261-266, 275. (in Chinese)
- [8] 胡云鸽, 苍岩, 乔玉龙. 基于改进实例分割算法的智能猪只盘点系统设计[J]. 农业工程学报, 2020, 36(19): 177-183. HU Yunge, CANG Yan, QIAO Yulong. Design of intelligent pig counting system based on improved instance segmentation algorithm[J]. Transactions of the CSAE, 2020, 36(19): 177-183. (in Chinese)
- [9] 高云, 郭继亮, 黎煊, 等. 基于深度学习的群猪图像实例分割方法[J]. 农业机械学报, 2019, 50(4): 179-187. GAO Yun, GUO Jiliang, LI Xuan, et al. Instance-level segmentation method for group pig images based on deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 179-187. (in Chinese)
- [10] MARTIN R, ACHIM K, FELIX A, et al. Automatically detecting pig position and posture by 2D camera imaging and deep learning[J]. Computers and Electronics in Agriculture, 2020, 174: 105391.
- [11] 刘岩, 孙龙清, 罗冰, 等. 基于改进 CNN 的多目标生猪检测算法[J]. 农业机械学报, 2019, 50(增刊1): 283-289. LIU Yan, SUN Longqing, LUO Bing, et al. Multi-target pigs detection algorithm based on improved CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 283-289. (in Chinese)
- [12] 高云, 李静, 余梅, 等. 基于多尺度感知的高密度猪只计数网络研究[J]. 农业机械学报, 2021, 52(9): 172-178. GAO Yun, LI Jing, YU Mei, et al. High-density pig counting net based on multi-scale aware[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 172-178. (in Chinese)
- [13] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.
- [14] 李菊霞, 李艳文, 牛帆, 等. 基于 YOLO v4 的猪只饮食行为检测方法[J]. 农业机械学报, 2021, 52(3): 251-256. LI Juxia, LI Yanwen, NIU Fan, et al. Pig diet behavior detection method based on YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 251-256. (in Chinese)
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C] // Proceedings of the European Conference on Computer Vision, 2018: 3-19.
- [16] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [17] YU Y, YANG X, LI J, et al. A cascade rotated anchor-aided detector for ship detection in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 60: 1-14.
- [18] ZHOU K, ZHANG Z, GAO C, et al. Rotated feature network for multiorientation object detection of remote-sensing images[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 18(1): 33-37.
- [19] ZHANG Z, GUO W, ZHU S, et al. Toward arbitrary-oriented ship detection with rotated region proposal and discrimination networks[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(11): 1745-1749.
- [20] YANG X, YAN J, MING Q, et al. Rethinking rotated object detection with Gaussian Wasserstein distance loss[C] // International Conference on Machine Learning, 2021: 11830-11841.
- [21] QIAO Y, MATTHEW T, SALAH S. Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming[J]. Computers and Electronics in Agriculture, 2019, 165: 104958.
- [22] ZOU Z X, SHI Z W, GUO Y H, et al. Object detection in 20 years: a survey[J]. arXiv preprint arXiv: 1905.05055, 2019.
- [23] LI S J, ABUFARHA Y, LIU Y, et al. Ms-tcn+: multi-stage temporal convolutional network for action segmentation[J/OL]. IEEE Transactions on Pattern Analysis and Machine Intelligence. DOI: 10.1109/TPAMI.2020.3021756.
- [24] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [25] GIRSHICK R. Fast R-CNN[C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [26] MUHAMMAD M B, YEASIN M. Eigen-CAM: class activation map using principal components[C] // 2020 International Joint Conference on Neural Networks, 2020: 1-7.